DOCUMENT RESUME

ED 123 886 FL 007 579

AUTHOR Jarvis, Gilbert
TITLE Testing for Meaning and Communication in the Foreign Language Classroom.
INSTITUTION Wisconsin Association of Foreign Language Teachers, Whitewater.
PUB DATE 75
NOTE 7p.
JOURNAL CIT WAFLT Bulletin; Fall 1975

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Communication (Thought Transfer); Communication Skills; *Diagnostic Tests; *Language Instruction; Language Skills; *Language Tests; Personal Growth; Role Perception; *Self Actualization; Self Concept; Semantics; *Test Construction; Testing; *Verbal Communication

ABSTRACT

Testing in foreign language classrooms is characterized by excessive preoccupation with students' ability to manipulate small grammatical features, while testing of communication is conspicuously absent. Furthermore, current testing is often done for the purpose of generating labels for students or for their post-instructional performance. This paper suggests that evaluators add another purpose: to discover what they know already and what they don't know yet. Testing of communication of meaningful content, with focus on all four language skills, should also be added. In this paper test items exemplifying these concepts are contrasted with traditional test questions. Although the incorporation of these concepts in foreign language testing will not radically change the status quo, perhaps it will promote the development of students who are more capable of communicating with other human beings, and help them to become more knowledgeable, sensitive, self-actualizing and fully functioning individuals. (DB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The WAFLT Bulletin, Fall, 1975

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Wisconsin Ass'n. of FL Teachers TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

ED123886

TESTING FOR MEANING AND COMMUNICATION IN THE FOREIGN LANGUAGE CLASSROOM

Gilbert Jarvis
Ohio State University

Testing is an extraordinarily vast topic. For example, at Ohio State we have a four hour course in foreign language testing alone. So we want to focus on just a few aspects of testing, aspects that influence, I believe, the self-concepts of our students. I want to create a background in terms of the roles our students play and how these roles contribute to self-concepts. I want to talk about the way we label our students through testing. I will end with some specific test items which show the distinction that I am postulating here as so important.

Each of our stdents learns to look at himself as he thinks others look at him. We, as teachers, are a significant part of the world, talking to the student, sending him messages about his worth; and often we do that through testing. In testing we end up labelling our students. We apply adjectives to them much more often that I would like us to, and, as I see it, at the wrong times. Sometimes we use a coding system for those adjectives. Instead of using the adjectives directly, we use symbols for them: "A", "B", "C", "F", or a number system. That is a very direct way to influence self-concepts. We say directly to the student, "You are great." "You are near great." "You are not so near-great." Sometimes we try, maybe out of an intuitive sense of guilt, to be a little less direct. We make distinctions and say to the student, "It is what you do, not what you are, that I am evaluating. And what you do is unsatisfactory, substandard, inferior." But the student hears, "You are unsatisfactory. You are substandard. You are inferior." It does not have to be that way. I know that the world out there sometimes expects us to label student behavior. They want our labels. In a sense, we work in a system that requires that kind of summative evaluation, evaluation that summarizes performance after the instruction has terminated. Also, we have done that for so long that we have a difficult time imagining anything else.

Testing can have other purposes than creating a label for someone. We can use it to make choices. We can use it to get information on how to help the student learn. What has he learned? And what has he not learned yet? Somehow we don't do this very often, the notable exception being in many individualized classrooms. All we are changing here is the purpose, not necessarily the test items. This formative or diagnostic evaluation occurs for the sole purpose of helping the student to learn--during instruction, not when it is too late to do anything about it. I'm making an appeal quite simply that we add, and that word is carefully chosen, not substitute, add some evaluation of this type. And that type of evaluation does in fact change the teacher-student relationship.

FL007579

There is a second way in which we send messages to students about their worth, messages which accumulate to become the way they see themselves. This way is far more subtle, and it does have implications for the kinds of test items that we create. I'm not going to talk about what is usually called "nonverbal behavior," but let's acknowledge that it is nonetheless real. Instead, I would like to talk about the roles that we and our students play or enact. The concepts of roles and role expectations are not at all exotic. We all play roles everyday. From moment to moment we move from role to role. If we were to open a social psychology textbook, we could read something that would say in essence, that in any given social system, there exists a set of expectations about how a person in any given position should act. These expectations define the role, and for better or worse we all seem to learn them rather quickly. I was myself struck by the significance of the way we enact roles just the other day when I took my car to the garage to have some work done. I was sitting there in the waiting room and I was suddenly overwhelmed with the sensation of being someone I didn't recognize stepping out of myself and asking, "Who's that?" I sat in the waiting room for awhile, paging through magazines that I probably wouldn't read otherwise. I got up and wandered very slowly, looking at things I would never otherwise look at, like the ceiling. I stopped and stared at an advertisement for maybe two minutes, the advertisement had I think six words on it. I walked very slowly into the showroom, look inside the new cars and at the window stickers, stepped back, kicked the tires. But if I were to do those things in a parking lot, someone would put me away. I acted that way because I was playing the role of someone waiting at an automobile agency for service. Not simply the role of a person waiting, for if I go to a doctor's office it's a different kind of waiting role. My point is that all life can be viewed in terms of moving from role to role. From one role to another, I function according to the expectations that we all have for a person occupying that position. When I go back to the garage in a few weeks, I will again play that role. My behavior will probably

resemble closely what it was the last time I was at the garage. But until then, I'll never do those things.

What is the significance of roles? As you or I or our students move from role to role, messages come to us about our success in the various roles. "You're a great teacher, Miss Jones." "You're a good football player, Scott." "You're a poor student, Cliff Evans." But in addition to these messages, the very nature of the role we create or impose matters a great deal. And that leads me directly into the foreign language classroom. I am unhappy with the role that we now impose on our students, impose by means of the expectations that we create in our classroom activities and our tests. The role we impose on our students is often a role that does not say to the student, "You are a real person whose ideas matter," or his foreign language behavior so often does not involve ideas. His activities and tasks deal less with meaning than with the mechanical manipulation of words. He may never come to see that words don't mean, people mean. And if your ideas don't seem to matter, then do you really count? The role of the foreign language student tells the student that he does not count in other ways. In class, what do we ask him to do? So often we ask him to listen, repeat, manipulate, translate. On our tests we ask him to find the one correct answer that we teachers have in our heads. How often do we say to him in one form or another, "Tell me what you think?" The role that we assign him is being judged continually. We evaluate what he does all of the time, either formally or informally. But scarcely ever do we ask him what he thinks or how he evaluates any topic, issue or idea. We don't ask him to make judgments, just to submit to them. And the message the student receives is, "Your ideas are not worthy of being heard. They are not important now." So I am asking for testing and teaching procedures that do three things. I am asking for testing and teaching procedures that, first of all, involve meaning and communication in any one or more of the four language skills. Second, I am asking for teaching and testing procedures that ask the student to express what he believes, thinks, and feels. Third, I am asking for testing activities and procedures that are structured to permit the student to express his ideas within the range of his linguistic competence. I am asking for procedures that are within the range of what he <u>can</u> do in terms of language, vocabulary, and grammar.

I think we all know what happens if you say to a first-year student, "Discuss such and such." How can we accomplish this goal? I do have a proposal that I think helps a little. As a general principle, let's stop always seeking one correct answer. Let's open-end our test items. Let's add some items. Notice I'm carefully using the word <u>add</u>, not saying we need to eliminate everything we're doing, or anything like that. Let's add some items that involve real meaning. Items where the student has to send or receive meaning. That's the way I define communication. Sending or receiving meaning in any of the four language skills. Let's create items that invite the student to express himself. Even if we have to compromise some of the conventional measurement values for what I believe is a higher value: self-expression.

Let's look at some typical test-item formats and ask how we could change them to accomplish these values. And I would like to begin with multiple-choice, because I think many of you would say that is the least humanistic type possible. I think we can modify that.

Here is a simple, straightforward multiple-choice comprehension question.

> Jaunquile likes being alone in his room. When his friends go to the show he prefers to stay home and listen to the radio or read. He is very:
> a. aggressive.
> b. solitary.
> c. extroverted.

Seeing just that much, in that rather standard test format for measuring listening or reading skills, you might ask, "What is new or different?" If the student sees it, it measures reading skill, if he hears it, it is listening. That kind of a format does involve meaning; the only way to answer it is to understand it. But I think the item becomes dramatically different by adding one element: the "d" alternative, i.e. "d?" The item now provides an option to be creative, an invitation to express any idea the student chooses.

Other items are similar.

> Pierre is always giving orders to his friends, to his brothers, and to his sisters. That indicates he has a tendency to be:
> a. patient.
> b. docile.
> c. authoritarian.
> d. invitation to say what you like.

I believe that "d?"makes a tremendous difference. It is an invitation to express any idea you want. Remember the subtle message I started with--an invitation to express your ideas says I am interested in your ideas. I am interested in the way you see these situations, these fictitious people even. I am interested in the way you view these people. I intentionally call the "d?" an invitation. Who can decline an invitation. If the student prefers the security of chosing the given, correct, or common, standard answer, it is there for him. If he doesn't know the keyed response, maybe he doesn't understand that one word, but if he understands everything in the stem, or question, part, it's rather realistic, because in language use we can always use another form if we're not sure of a word.

Of course, there is a small flaw in the term: the item adds a productive skill, presumably writing. And that means the item does not have the same kind of purity of measurement, it does not measure reading skill alone. It's also less simple to score now. But at what price objectivity? I'd rather do without it in a lot of places. In order to evaluate, in order to score an item such as this, you have to look not only at the keyed acceptable answer, but any time a "d" is used, you have to evalute that response also. That's a little bit more work. Also, there are probably going to be a few instances where a student ventures his own answer and is wrong. when he could have chosen a secure alternative. And when that happens I am bothered, but fortunately that is a weakness which in practice seems to occur only rarely.

What are the benefits? First, there is a total focus on meaning. There is no mere manipulation of pronouns, verb tenses, or adjectival agreement. No "planned parrot-hood," as someone called it. Using such items, items that focus on meaning both in testing and in classroom practice, creates a set or expectation for meaning, for the sending and receiving of meaning. And isn't that what language is for? In a sense, we could view all of our lives as an effort to share our thoughts and feelings with someone else. Let's consider one more example of the multiple-choice format. This one is much more personal.

When I read a book I:
a. like to drink Coke.
b. keep my eyes closed.
c. make the room totally dark.
d. whatever you'd like to say.

I am usually enthusiastic when:
a. I go to basketball games.
b. a good friend is ill.
c. I loose something valuable.
d. whatever you'd like to say.

There is, I think, something to keep in mind whenever classroom practices or test items become personal. I get very uncomfortable with some of the humanistic or affective education techniques that are in vogue these days. There is a point at which they start to probe into one's innermost soul and get close to something like psycho-therapy, and that's a problem. In any group of 25 people there are some who are uncomfortable even with elements that are less personal than that. And I'm afraid some of the practitioners often overlook that aspect of their humanistic of affective tendencies. I think that is something to keep in mind here. If, for example, we provided a student with only four choices, all rather personal, we might be asking him to say something he may not want to say. When you provide an out, the question mark, he can be as personal or as impersonal as he wants to be. I think that matters.

I have an option for those of you who are uncomfortable with any of these mechanisms that we're talking about as testing devices. Try them for classroom activities, for practice activities, because they are still going to lead to the goal I have in mind for foreign language education.

Let's look at another mechanism, one we can modify from our conventional understanding of it, our conventional stereotype. It might be called matching. The minute I say the word matching, we tend to think of one column of items where every item links to an item in another column. Well, matching here has a different meaning. By matching, I mean combining elements to say something you want to say. Here's one with an ecological base. "If you are concerned about environmental protection, you may some day want to make signs for demonstrating or for advertising your concern. Write six ecological slogans by combining different items from each column, or by adding your own items." So the first column has perhaps imperative verb forms--"down with, abolish, eliminate, free, limit, support, no more, save," and the invitation. The second column has items such as, "small cars, trees and flowers, superhighways, air pollution, clean lakes, large cars, detergents, the fish in our streams, clean air." The student puts items together like, "Down with air

pollution." "Limit large cars." "No more superhighways." That's the ultimate in linguistic simplicity. Why? Because all of the elements that the student needs are there for him. He just has to put them together. But what does it require to put them together? Meaning. Therefore, this particular type can be used right from the beginning in a first year class. In fact all of these items are variations, they are adaptations from learning activities that a team of people at Ohio State and I put together in a little elementary French reader that we are intending for second semester college or second year high school. It is my conviction that this combining or matching mechanism can be used right from the beginning. And I am not uncomfortable at all with a section on a test that's very thematic, such as an ecology section . In fact, I am much more comfortable with that than with a section that says "relative pronouns." In other words, a semantic base rather than a grammatical base. These are open-ended both in the sense that many right answers are provided and there is also that invitation to express whatever else the student chooses and can express. It is an invitation to be declined, if the security is more comfortable.

Let's look at another example, one that we might call cliches. Using an item from each column to form sentences, try to recreate the most common cliches about each of the different types of drivers. Write ten cliches. "Old people," and then from the other column, "tend to drive more slowly," is a most common choice. "Truck drivers think they own the road." "Italians..." I won't finish it, you do it. "Men, or women, drivers." The focus is on meaning. This one is thematically linked to a text. I created this for a text in this reader that is a driving test. It's called "Can You Drive in Europe?" and the basic reading is a driving test with European road signs and whatnot. So it is thematically linked; and that is an important characteristic. It focuses on semantic, rather than grammatical, structures. I am saying we need to add some items of this kind. I don't say we can't test a relative pronoun.

Now, let's look at one more in this format. Here is one that focuses on a grammatical concept. But contrast it with the kinds of things that are done in almost all textbooks. It's "make," and then you put in a number if you want, "X number of original sentences by combining elements from the two columns." Then you list the names of famous people from either the sports world or music world: "Billy Jean King, Hank Aaron, Louis Armstrong, Van Cliburn plays the..." and then a list of musical instruments or sports. How does this contrast with what we typically do? We have some good evidence that when all the student has to do is fill in the blank with the right preposition, he does not read for meaning or listen for meaning or whatever the skill happens to be. He simply looks at the overall task and says "What is the minimum information I need in order to complete this task?" And he plugs in what he needs. Here he has to get the meaning or he can't make the sentence. And I think that's a dramatic difference. There is still the invitation for him to say or add any figure, any person he wants to, because we teachers can't think of all possibilities, course. And he can add any sport or instrument that he wants.

How about another typical mechanism that we use? How can we modify it a little to develop a self-examining, self-aware communicator? Someone who has ideas that matter. Let's look at completion items. These tend to be more complex, and one has to make sure the student can do them. There has to be some prerequisite ability. There are many correct answers. And the subtle message is, subtle but real, that the student's thoughts matter. Such items can also help the student to examine his attitudes, beliefs, and feelings, without putting him on the spot.

Here's a second kind of item, creating a commercial for a real imagined product. It's a creative task, starting with "I want to speak to you about something you're going to like. We have a new product that is called _____. It is very good because _____. So if you want to _____, you ought to buy _____." There is flexibility. A student can complete this with one or two or three words; if he is capable, he can write a paragraph for any of those completions. Flexibility and options are built in it.

Here's another one with grammar. "Complete the following sentences with the appropriate country or city and be sure to use the correct preposition in French." Think about the kinds of things we put on our tests, or the activities like this in our textbooks. What happens? There is a blank in the middle, and all the student does is look at what comes after, and he puts in the preposition. He goes down the column rather than reading the sentences. No meaning involved in that. The whole exercise is a problem to be solved, and he says, "What is the minimal information I need to solve the problem?" That's all he does. But here, what does he have to do? He has to get the meaning in order to complete the sentence. "To see the Eifel Tower you have to go..." You don't know where to go, if you don't have meaning in there. "To France, to Paris." And you have to use a different preposition depending upon whether you chose Paris or whether you chose France. Meaning has to be involved. Next is an item that deasl with an overall theme. Everybody is concerned with an ideal life. "An ideal father is a father who..." In a unit with many

adjectives the student uses these, or he can use an entire sentence. "An ideal friend is a friend who..." "An ideal wife is a wife who..." "An ideal teacher is a teacher who..." It can be revealing.

The next category I call simulation. In this item we ask the student to function as if he were in some given situation or as if he were something or someone, to play a role in a sense. This one again has a thematic link to the unit. The unit of this reader which I referred to is "What Do You Do If You See a Flying Saucer?" That's the reading passage. Then as a test question, or as a learning activity, if you like to use them in class, this item continues: "A little green man from another planet has observed some strange phenomena during his visit to Earth. He asks an inhabitant of Earth to help him identify these different phenomena. Can you help him?" Number one is, "In a large city in the north of France I saw a very strange metallic tower, very tall, and there are always many people around this structure. They take pictures. Often these people get into a little box that goes up and down continually. What is this weird structure?" Total comprehension is needed in order to answer. The student may guess from the second sentence on, "Oh, it's the Eifel Tower." But what happens in practice? We find that they read every sentence, trying to confirm their hypothesis, to make sure they are correct. And that is seeking meaning. A second example: The little green man says, "Can you explain to me the bizarre scene that I observed several times? There were men running in all directions, some dressed in red, others in blue. They kick a round object with their feet. There are many spectators who shout and jump a lot. What is this?" There is a total focus on meaning here. This item is not open-ended. But at a high enough level, second year probably, my next task with these students, my next test item or teaching activity, would be to have them create a similar situation in American culture. Either alone or in small groups. The create a description of something and then work with a group of students and have them guess what it is. That's a very productive, creative task. Here is another example. "You've invited four friends to have dinner with you. You go to the market to buy food. You buy meat, green beans, etc. How much will you buy?" This is related to a unit called "Metric Shock." A fellow by the name of Miles Long wakes up one morning and the whole world has gone metric. (He puts on his sweater because he has the radio on and he gets the temperature in centigrade; jumps into his car and gets a ticket right away because the speed limits are different.) This is a simulation activity that measures comprehension. The second part says, "Now the meat costs 30 francs per kilo, potatoes 1.9 francs per kilo, oranges, etc. How much money did you spend?" Again, total focus on meaning.

Let's see what we can do with the traditional true-false format. Conventionally that has meant comprehension of some text or content. And I happen to believe that that's superior to those manipulative pattern drills. This is a variation of the true-false format. We start with a car, and the student must indicate whether the item is true or false and then change any false statement to true. The first one says, "I'm going to put my baggage in the trunk." Number three says, "When you're driving, keep both hands on the gear shift." Another says, "Before backing up, look carefully in the glove compartment." The visual is not here just by accident. If you want to add vocabulary, there's a place to add it, just label parts of the car. There is not a great deal of freedom of expression as this is structured, but there is a little, because the student can choose how he wants to make it a true statement. He can change it in lots of different ways. But the focus is on meaning. When I use items like this, I would put in more of the false, or negative alternatives than the true, perhaps 75% or so. Let's change that format a little bit, to something that is definitely less precise but which does focus more attention on the student than on verbs. "Indicate whether you agree or disagree with the following sentences. Correct those you disagree with or which are not true for you. 'At home I have to fix all the meals for my family.' 'At home my mother fixes all of the meals for the family.' 'At home I do not have to fix all the meals for my family.'" If you are uncomfortable with the fact that some students may just put "nots" in and take the easiest way out, then underline some part of the sentence in all of the responses and tell the student to change the underlined part if he wants to make it a true statement. "I like to sew every weekend." "Women should work at home instead of having careers." With such items, students will express values very vividly. You're inviting it. Words here are for communicating, not for pushing around like little colored blocks. The continual message, sent even by the test is that the student's ideas matter. These statements are here for comparison with student ideas and values. Therefore, I'd rather call these items "alternate choice with correction" rather than true-false. Students can agree, disagree, the choices might be plausible, implausible, applicable to you or not applicable to you.

How about an example for the speaking skill? My example for the speaking skill also focuses on communication, and communication always involves more than one person. So I'm going to have my students cooperate often when I measure speaking and listening. For this activity, one student has a visual, maybe a little chalkboard with items drawn on it, or a magnetic board, or maybe just a

picture, photograph, or line drawing. And the other student faces him so that they cannot see each others boards. The presenter has a picture already on his board, and his task is to describe it so that the receiver can draw a presentation of it or can place the objects on the board in the right place. Some teachers don't permit gestures, some do. Usually the teacher assigns a rating both to the presenter and to the receiver. The teacher might choose to use a "can do/cannot do/cannot do yet" evaluation instead of a label. Real communication.

As a last item, have you ever tried a test for fun, a self-test? It's not a test, it's a fun activity, but it looks like a test, so I thought I'd mention it. Here is one. "Can you keep your cool? If you want to find out, take the following test. The situations described denote varying degrees of fear in different people. Use the numbers one to five to indicate your reaction and write the number beside each sentence. One is, "Petrified with fear," and the other extreme is, "Don't even pay attention to the situation." Sample items are: "It's night time, you are alone, you hear a strange noise in another part of the house." A student indicates on a scale from one to five how afraid he is. "You're in a plane over the Atlantic, and the four jet engines quit." "You're going to be married in 15 minutes, or the ceremony will begin in 15 minutes." "You stop at a service station to use the rest room, as you go to leave, the door will not open, you shout but the station is closed for the night." So the student marks from one to five and you then ask him to add up his score and divide by six, or however many items you have. Then you give him a humorous interpretation. Tell him whehter he can keep his cool or not. It's all for fun. If his average is one to one and one-half, the interpretation is, "Be careful, you're in danger of a heart attack." If it's one and one-half to two, "You're very alert to danger but unfortunately your self-control is inadequate." The next two are relatively innocuous, then four and one-half to five is, "It's not coolness, it's apathy."

Let me try to synthesize some of the points I've discussed. I've tried to emphasize, and maybe even overstate, a point of view that holds that testing in our foreign language classrooms is characterized by excessive preoccupation with the student's ability to manipulate small grammatical features. Testing of communication is conspicuously absent. Furthermore, we often test for the purpose of generating a label for the student or for his performance after the instruction has terminated. I have not advocated anything radical, like throwing out all of the grammatical or all of the summative evaluation. I have carefully used the word add. I'm suggesting that we add another purpose. The purpose is to make some of our testing discover how to help a student learn. How do we discover what he knows and what he doesn't know yet? I also want us to add some testing of communication dealing with meaning and all of the language skills. By incorporating some of these concepts in our testing, we won't radically change the nature of foreign language education or the course of society. But maybe we will develop students who will communicate just a little bit better with parents, friends, lovers, bosses, with anyone. I want to create a role that helps the student come a little bit closer to what Masslow calls the self-actualizing person, to what Rogers calls a fully-functioning person, to what any of us would call a more knowledgeable and sensitive person.